



# Transmission robuste de vidéo basée ondelette à travers un canal MIMO

Julien Abot, Clency Perrine, Christian Olivier, Yannis Pousset

## ► To cite this version:

Julien Abot, Clency Perrine, Christian Olivier, Yannis Pousset. Transmission robuste de vidéo basée ondelette à travers un canal MIMO. CORESA 2010, Oct 2010, Lyon, France. hal-00624517

**HAL Id: hal-00624517**

**<https://hal.science/hal-00624517>**

Submitted on 19 Sep 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Transmission robuste de vidéo basée ondelette à travers un canal MIMO

J. Abot<sup>1</sup> C. Perrine<sup>1</sup> C. Olivier<sup>1</sup> Y. Pousset<sup>1</sup>

<sup>1</sup> XLIM-SIC (Laboratoire XLIM département Signal Image Communication)

CNRS UMR 6172, Université de Poitiers,  
Téléport 2, Bvd P. et M. Curie, BP 30179, 86962 Futuroscope-Chasseneuil Cedex – France

{abot, perrine, olivier, pousset}@sic.univ-poitiers.fr

## Résumé

*Cet article présente l'évolution d'un codeur pour images fixes basé ondelette vers le codage vidéo. Ce codeur associe une transformation en ondelette classique, une quantification vectorielle par sous-bande et une modulation de type MAQ-M. L'exploitation conjointe des éléments de la chaîne de transmission assure une qualité de service dans des conditions de transmission difficiles. Les redondances temporelles sont exploitées, dans le cadre du codage vidéo, par un algorithme spécifique d'estimation et de compensation de mouvement adapté à nos contraintes de codage. Nous proposons ainsi une méthode de codage permettant d'atteindre une bonne robustesse des vecteurs mouvement. Ce codeur nommé V-WTSOM présente une hiérarchisation importante des informations. Nous exploitons cette hiérarchisation en transmission à travers un canal MIMO incluant une possibilité de précodage diagonal. La décomposition en sous-canaux SISO parallèles et indépendants permet d'envisager des stratégies d'allocation basées sur le contenu et améliorant globalement la qualité de service.*

## Mots clefs

Transformée en ondelette, vecteur mouvement, MIMO, précodage, V-WTSOM

## 1 Introduction

De nos jours, nous assistons à un développement important des technologies sans fil telles que la téléphonie mobile ou les réseaux ad'hoc. Ces technologies impliquent des transmissions d'informations dont la qualité et la quantité ne cessent d'augmenter à l'image de la HD. Toutefois la bande passante sur les canaux radiomobiles devient de plus en plus limitée. On comprend alors l'importance que revêt l'étape de compression dans une chaîne de transmission vidéo. A l'heure actuelle, de nombreux standards de codage vidéo (H.264, MPEG-4) permettent d'obtenir un compromis intéressant en termes de débit/distorsion sur des canaux présentant peu de perturbations [1]. Tous les standards vidéos actuels sont basés sur la transformée DCT

(Discrete Cosines Transform). Aujourd'hui, aucun codeur basé ondelette n'a fait l'objet d'une standardisation. Pourtant différentes techniques de codage vidéo basées ondelettes existent dont [2, 3] sont des exemples. L'intérêt de ces méthodes est de compresser davantage en apportant une scalabilité accrue. Mais ces méthodes s'occupent rarement des problèmes de robustesse lors de la transmission. Comme source de problème, nous pourrions citer, par exemple, le codage à longueur variable qui peut être responsable de désynchronisations de flux en réception. Des méthodes de robustesse [4, 5] ont été envisagées pour H.264 donnant naissance au profil « extended ». Toutefois lorsque le Taux d'Erreur Binaire (TEB) devient trop important il devient nécessaire d'utiliser des Codes Correcteurs d'Erreur (CCE) ayant pour conséquence de diminuer le débit utile. Afin de faire face à cette problématique, nous avons proposé dans [6] une stratégie de codage conjoint (codeur WTSOM pour Wavelet Transform Self-Organized Map pour images fixes). Cette stratégie a été pensée dès la conception du codeur pour s'adapter aux contraintes apportées par le canal de transmission et non les subir. Ainsi, il devient possible de transmettre des images sur des canaux présentant des TEB très importants ( $\geq 10^{-3}$ ) et cela, sans CCE. Le codeur vidéo V-WTSOM présenté dans ce papier, s'appuie comme pour les standards sur l'estimation et la compensation de mouvement. Il exploite aussi les qualités du codeur WTSOM. Pour les codeurs standards, les Vecteurs Mouvement (VM) sont très sensibles aux erreurs de transmission. Nous proposons pour V-WTSOM une technique d'estimation et de codage des VM permettant d'assurer la robustesse au prix de certaines distorsions. L'objectif étant de décoder une vidéo dans des conditions où les standards n'arrivent plus à décoder. Par exemple, V-WTSOM présente un intérêt pour la visioconférence. L'objectif de ce codeur n'est pas de concurrencer les standards mais d'assurer des transmissions exploitables lorsque les standards ne peuvent plus assurer la qualité de service requise en minimisant le débit utile de la source. Les conditions de transmission défavorables se retrouvent en particulier dans les environnements complexes tels que les environnements urbains riches avec multitrajets et/ou

incluant la mobilité. La technologie MIMO (Multiple Input Multiple Output) couplée à la modulation OFDM (Orthogonal Frequency Division Multiplex) permet d'exploiter la diversité du canal. Cette association a déjà été adoptée pour les normes Wi-Fi (IEEE802.11n), LTE et Wi-Max (802.11e). Nous utilisons cette association dans le cadre de la stratégie de transmission V-WTSOM afin d'exploiter la hiérarchisation du flux de données. L'apport de ce papier est donc de proposer une stratégie de codage robuste des VM permettant une bonne adéquation avec une stratégie de transmission basée MIMO.

Dans la seconde partie, nous présenterons le codeur d'images fixes WTSOM ainsi que son extension à la vidéo. Dans une troisième partie, nous évoquerons la technique MIMO et en particulier le principe du précodage diagonal et son intérêt pour la stratégie de transmission. La quatrième partie présentera des résultats de simulation avant de terminer par une conclusion.

## 2 Codeur V-WTSOM

### 2.1 Codage en GOP

Le codeur V-WTSOM exploite les redondances temporelles par le codage en GOP (Group Of Pictures), à l'image des standards. On trouve 3 types d'images : Intra (I), Prédite (P) et Unidirectionnelle (U). Le codage de ces images obéit à une hiérarchie illustrée sur le schéma suivant (fig. 1) :

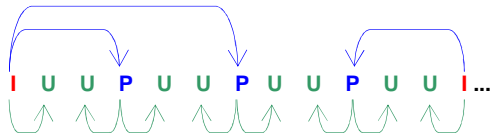


Figure 1 - Codage du GOP dans V-WTSOM

Sur la figure 1, la pointe d'une flèche indique une image courante. Ces images nécessitent une image référence (désignée par le départ d'une flèche) pour leur reconstruction. Les images Unidirectionnelles sont des images nécessitant une seule référence afin d'être reconstruite, à la différence des images Bidirectionnelles des standards, qui nécessitent une interpolation liée à deux références.

### 2.2 Codage des images Intra

Les images Intra sont codées suivant l'algorithme WTSOM pour images fixes [6] dont nous rappelons brièvement le principe (fig. 2). L'image subit une transformation en ondelette de niveau 3. Seules les sous-bandes les plus informatives sont conservées (5 à 7 suivant la QoS désirée). Pour justifier ce choix, nous nous appuyons sur l'entropie calculée dans chaque sous-bande [6]. L'impact visuel lié à cette suppression d'information est jugée acceptable compte tenu de l'application visée. Les sous-bandes conservées subissent une Quantification

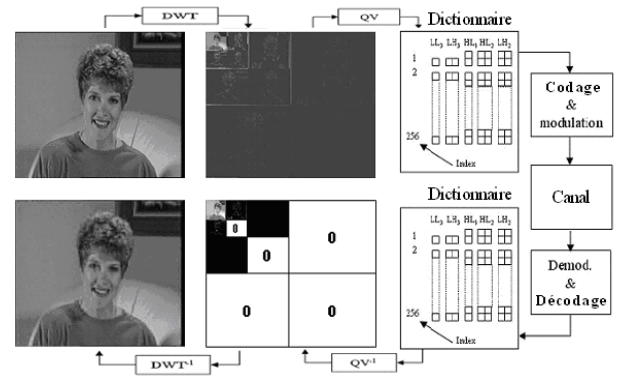


Figure 2 - Principe de codage de WTSOM fixe

Vectorielle (QV) dont le format des vecteurs de quantification est adapté au contenu de la sous-bande. Réaliser une QV implique d'utiliser des dictionnaires de vecteurs. Pour les construire, nous utilisons les cartes topologiques de Kohonen issues de l'algorithme Self-Organizing Map (SOM). Cet algorithme présente la particularité de générer des dictionnaires auto-organisés dans le sens où deux vecteurs voisins dans le dictionnaire vont présenter le plus de similitudes possibles. La stratégie WTSOM tire sa robustesse de cette particularité. En superposant une constellation MAQ-M aux dictionnaires auto-organisés, nous pouvons minimiser l'impact d'une erreur de transmission (en considérant M égal au nombre de vecteurs du dictionnaire). Dans cette configuration, un symbole erroné en réception produit un vecteur décodé très proche du vecteur original. La distance euclidienne sera d'autant plus grande que l'erreur sera importante. Cette stratégie permet d'atteindre de bonnes performances en termes de robustesse. Dans le cadre d'une transmission, nous considérons que le récepteur dispose des dictionnaires. Le débit atteint pour un codage type [6] est de 0,3125 bpp.

### 2.3 Codage des images P et U

Le codage des images P et U s'inspire des techniques employées par les standards MPEG et H.26X. Le codage de ces images est réalisé en exploitant les redondances temporelles par le biais d'un algorithme d'estimation et de compensation de mouvement par block-matching (BMA pour Block-Matching Algorithm) [7].

Rappelons que l'estimation de mouvement par BMA implique d'utiliser une image référence et une image courante. L'image courante est découpée en blocs, l'idée consiste alors à rechercher où les blocs de l'image courante se situent dans l'image référence. Un bloc est désigné comme étant optimal par minimisation d'un critère, généralement le SAD (Sum of Absolute Difference). Le déplacement des blocs est représenté par un VM créant ainsi un champ de VM pour l'image complète. Appliquer ce champ de VM sur l'image référence permet de reconstruire une image compensée en

mouvement. Toutefois, un champ de VM ne peut pas coder à lui seul une image car celle-ci présenterait de multiples distorsions (effet bloc important). Ainsi un terme d'erreur (résidu) entre l'image courante originale et l'image compensée en mouvement est calculé. Ce résidu doit subir un codage intra avant sa transmission.

### 2.3.1 Estimation de mouvement

Dans le standard H.264, le champ de VM a pour but de donner une estimation de l'image (image compensée en mouvement). Un résidu va contenir les corrections à apporter à l'image. En outre, l'estimation et la compensation de mouvement par le champ de VM permettent de réduire l'entropie du résidu. Par la suite, le résidu est compressé (découpage en blocs, DCT, quantification, codage à longueur variable). La faible entropie présentée par le résidu couplée à un codage à longueur variable permet ainsi d'être efficace d'un point de vue compression. Par contre la méthode est sensible aux erreurs de transmission (désynchronisation du flux). Il faut alors mettre en place des outils de robustesse comme par exemple des CCE ou des marqueurs de resynchronisation. Pour V-WTSOM, l'approche est différente. L'objectif de ce codeur est d'être robuste aux erreurs de transmission. Ainsi, nous avons opté pour un codage à longueur fixe (absence de désynchronisation du flux en cas d'erreur). Cette approche rend obsolète la notion d'entropie du résidu car la quantité d'information, faible ou importante, n'aura pas d'impact sur le taux de compression des résidus. La principale difficulté consiste alors à réaliser un codage efficace mais respectant la contrainte suivante sur les taux de compression au sein du GOP, i.e :

$$T_c(I) < T_c(P) \leq T_c(U)$$

où  $T_c$  est le taux de compression des images I, P ou U.

### 2.3.2 Codage des VM

La problématique du codage des VM est récurrente dans le codage vidéo. Les VM peuvent être vus comme des pointeurs vers une zone de l'image référence. En complément le résidu permet de corriger localement la distorsion éventuellement présente lors de la compensation de mouvement. On comprend alors que la moindre erreur de transmission sur les VM engendre un pointeur vers une autre zone de l'image référence. Cette zone n'est alors plus optimale et le résidu devient inefficace afin de corriger la distorsion. Les VM sont, par conséquent, très sensibles aux erreurs de transmission qui modifient leur valeur. Dans le standard H.264, les VM subissent une prédiction spatiale, puis sont codés sans perte à l'aide d'un codeur entropique (CABAC). Dans le schéma de codage V-WTSOM, cette sensibilité aux erreurs doit être maîtrisée afin d'assurer la robustesse. Afin d'être en accord avec le schéma de WTSOM, nous

souhaitons que notre codage soit adapté à une transmission MAQ-M. Une étude a permis de conclure qu'un codage sur 4 bits des VM représente un bon compromis entre qualité/distorsion et robustesse. L'algorithme BMA va venir lire le contenu du dictionnaire, et ne va tester que les blocs candidats désignés par les VM du dictionnaire. Se pose alors le problème de la construction de ce dictionnaire. Utiliser une carte topologique de Kohonen pour la construction de ce dictionnaire ne donne pas de résultats cohérents relativement à une mesure de qualité. Notre proposition est basée sur une simple étude statistique des VM (fig. 3) sur une base de sept vidéos.

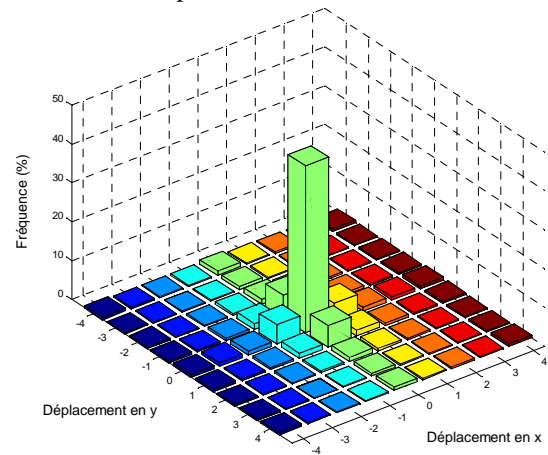


Figure 3 - Distribution des VM après estimation exhaustive sur un ensemble de sept vidéos

La figure 3 montre la distribution des VM lorsque l'estimation se fait de manière exhaustive (tous les blocs de la fenêtre de recherche sont testés). Le VM (0,0) présente une fréquence plus importante que les autres VM. Il correspond à tous les blocs qui ne bougent pas d'une image à l'autre. Les VM les plus courants sont aussi dans le voisinage de (0,0). Ainsi les 16 VM les plus fréquents représentent à eux seuls 77,3% de tous les VM d'une estimation exhaustive soit une bonne représentation. La concentration des VM les plus fréquents autour de (0,0) impliquera qu'en cas d'erreur de transmission, l'impact visuel sera minimisé par le choix d'un bloc très voisin du bloc optimal. Enfin, considérer un nombre de VM supérieur à 16 reviendrait à rendre notre algorithme plus sensible aux erreurs de transmission. En effet, plus de VM implique des erreurs d'impact potentiellement plus important à cause de la distance euclidienne plus grande entre deux extrémités du dictionnaire. Finalement, du point de vue de la transmission des VM, il est nécessaire d'organiser le dictionnaire de manière à respecter l'ordonnancement spatial des VM en cas d'erreur.

### 2.3.3 Traitement des résidus

Le résidu bénéficie d'un codage Intra adapté à ces caractéristiques. Etant donné que, de part sa nature, il

contient moins d'information qu'une image Intra classique, nous pouvons envisager de le coder avec des vecteurs présentant une résolution plus importante. De plus, suivant la qualité désirée nous pouvons coder plus ou moins de sous-bandes. Afin d'obtenir des résultats acceptables, l'expérience montre qu'il n'est pas nécessaire de coder plus de 3 sous-bandes (LL3, HL3 et LH3) pour les résidus des images P et seulement LL3 pour les résidus des images U. Le codage et l'utilisation du résidu permet au codeur V-WTSOM de faire preuve d'une grande scalabilité.

## 2.4 Exemple de transmission

Nous présentons un exemple de transmission réalisée par le biais d'un canal gaussien et d'une modulation MAQ-16. La passage de la MAQ-256 à la MAQ-16 pour la transmission des éléments codés sur 8 bits est compatible mais provoque une baisse de PSNR de l'ordre de 1,5dB [6]. Les vidéos utilisées sont au format CIF en niveaux de gris. Les caractéristiques du codage V-WTSOM pour ces simulations sont données dans le tableau 1 :

Codage I	Codage P		Codage U		Tc
	VM_P	Résidu_P	VM_U	Résidu_U	
LL3 : vecteurs 1x1 HL3 : vecteurs 2x1 LH3 : vecteurs 1x2 HH3 : vecteurs 2x2 HL2 : vecteurs 4x4 LH2 : vecteurs 4x4 HH2 : vecteurs 4x4	Bloc BMA : 16x16 pixels	LL3 : vecteurs 1x1 HL3 : vecteurs 2x2 LH3 : vecteurs 2x2	Bloc BMA : 16x16 pixels	LL3 : vecteurs 2x2	0,114 bpp
Information codée sur 8 bits		Information codée sur 4 bits			

Tableau 1 - Exemple de configuration de codage pour le codeur V-WTSOM

Un des apports de ce travail réside dans la robustesse des VM au bruit de transmission, sans protection particulière. La figure 4 est un exemple de transmission permettant de mettre en avant la robustesse des VM. Dans cet exemple, seuls les flux codant les VM ont été bruités (fig. 4).



Figure 4 - Illustration de la robustesse des VM (Foreman\_40) ; (a) TEB des VM = 0 ; PSNR moyen = 25,4dB ; (b) TEB =  $1,4 \cdot 10^{-1}$  ; PSNR moyen = 23,7dB ;  $T_c = 0,114$  bpp

On remarque que de légères distorsions apparaissent. Elles s'expliquent par le fait que le bruit peut changer la valeur d'un VM. Cela signifie qu'au lieu de choisir le bloc optimal pour décoder une zone de l'image, l'algorithme sélectionne un bloc dans son voisinage immédiat

(relativement à l'importance du bruit). On peut espérer que localement les différences entre blocs sont minimales ce qui a pour effet de minimiser l'impact d'une erreur de transmission.

## 3 Stratégie de transmission MIMO

### 3.1 Principe du MIMO

Face aux contraintes propres du canal de transmission telles que la mobilité ou les phénomènes de multitrajets, des solutions multi-antennaires ont été développées. Le principe de base des systèmes multi-antennes baptisés MIMO (Multiple Input Multiple Output) [8] consiste à exploiter au maximum la diversité de transmission qu'elle soit spatiale, temporelle ou fréquentielle (fig. 5).

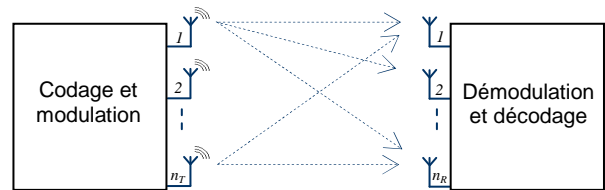


Figure 5 - Principe d'un système MIMO

Ces systèmes permettent d'envisager deux stratégies. On peut optimiser le débit en décomposant le flux à transmettre en  $n_T$  sous-flux transmis simultanément. On parle alors de démultiplexage spatial (LST pour Layered Space Time). On augmente ainsi le débit d'un facteur  $n_T$ . On peut également favoriser la robustesse de la transmission en transmettant  $n_T$  fois le flux originale. On parle alors de Codage Spatio-Temporel (CST). Dans ce cas de figure, le débit correspond au débit d'un lien SISO mais la robustesse du lien est améliorée. Des solutions existent également permettant de mixer ces deux approches.

### 3.2 Utilisation de précodeurs MIMO

Il existe beaucoup de configurations de systèmes MIMO. Ils peuvent être classés en deux familles suivant que l'on connaisse l'état du canal à l'émission ou à la réception. Cette connaissance est obtenue par le biais de séquence d'apprentissage. Si l'on dispose uniquement de l'état du canal à la réception, on parlera de système en boucle ouverte (OL-MIMO pour Open Loop MIMO). Dans ce cas, nous ne pouvons envisager qu'une stratégie de transmission aveugle. La solution la plus efficace consiste alors à transmettre avec une puissance équivalente sur chacune des antennes émettrices. Au contraire, si l'on a connaissance de l'état du canal à l'émission, on parle de système MIMO en boucle fermée (CL-MIMO pour Closed Loop MIMO). Nous pouvons alors envisager d'adapter la puissance d'émission sur chaque antenne afin d'optimiser la transmission de l'information. C'est le rôle des précodeurs diagonaux MIMO. Connaissant l'état du canal, les précodeurs peuvent décomposer le canal MIMO



en sous-canaux virtuels SISO (Single Input Single Output). Ces sous-canaux parallèles et indépendants présentent alors des caractéristiques propres (capacité, TEB, SNR...). Les précodeurs cherchent à optimiser la transmission de l'information suivant un critère et fournissent ainsi une pondération sur chacune des voies virtuelles SISO. L'utilisation de précodeurs diagonaux permet ainsi d'envisager une allocation intelligente de l'information et se prête donc particulièrement bien à la transmission d'un flux hiérarchisé.

### 3.3 Stratégie d'allocation V-WTSOM/MIMO

Le flux issu du codeur V-WTSOM est un flux hiérarchisé. En effet, il existe une hiérarchie entre les images issues du codage en GOP et également une hiérarchie entre les sous-bandes d'ondelette. Nous considérons dans cet article un système MIMO  $4 \times 4$  (quatre antennes à l'émission et à la réception). Nous disposons alors de quatre voies pour transmettre notre flux V-WTSOM. La stratégie d'allocation consiste à décomposer notre flux en couches de qualité hiérarchisées. Si le canal est bon, nous aurons donc accès à toute l'information et donc à la qualité maximale, alors que si le canal est dégradé, nous favoriserons une qualité de base assurée par la transmission d'une seule couche. Le schéma ci-dessous illustre la décomposition du flux à transmettre et leur allocation sur les sous-canaux SISO (fig. 6) :

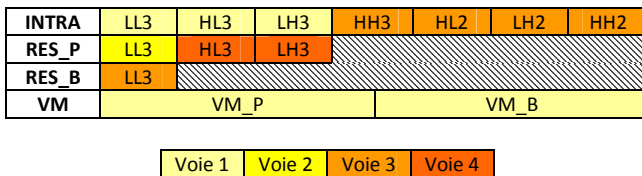


Figure 6 - Allocation du flux V-WTSOM sur les différentes voies virtuelles SISO

Ce découpage a été réalisé pour obtenir quatre sous-flux de taille équivalente correspondant à des couches de qualité successives. Pour décoder la couche  $n$ , il est nécessaire de disposer de toutes les couches inférieures à  $n$ . Ainsi la couche 1 constitue la couche de base indépendante des autres. Toutefois, les caractéristiques propres à V-WTSOM imposent d'obtenir les informations des voies 1 et 2 pour obtenir une vidéo reconstruite d'une fluidité optimale. Les voies 3 et 4 représentent ainsi des couches de qualité supérieure exploitables lorsque le canal est jugé bon.

## 4 Résultat de simulation

### 4.1 Chaîne de simulation

Les simulations ont été réalisées par le biais de la chaîne de transmission présentée sur la figure 7. Elle est initialement utilisée avec le codeur H.264/SVC et détaillée

dans [8]. Le principal changement par rapport à la chaîne originale concerne la suppression des CCE. La vidéo originale est codée suivant l'algorithme V-WTSOM avec les paramètres correspondant au tableau 1. Ce codage fournit un flux hiérarchisé et découpé en 4 sous-flux comme indiqué sur la figure 6. Le précodeur utilisé est WF (Water-Filling). Pour la phase de précodage, nous supposons une connaissance parfaite de l'état du canal (Tx-CSI (fig. 7) pour Transmitter Channel State Information). La modulation MAQ-16 est couplée à la technique de modulation multi-porteuse OFDM dont les paramètres, séquences pilotes et d'apprentissage correspondent à la norme 802.11n [9]. Du côté du récepteur, le décodage est assuré par un décodeur à Maximum de Vraisemblance (MV) avant démodulation OFDM et MAQ-16. Le canal de transmission est un canal sans mémoire dont les coefficients sont tirés suivant une loi de Rayleigh. De plus la transmission est perturbée par un bruit blanc additif gaussien de moyenne nulle et de variance égale à 1.

### 4.2 Intérêt de l'allocation MIMO

La figure 8 permet de constater l'apport en qualité d'une allocation MIMO utilisant le précodeur WF en comparaison de la stratégie SISO classique. Ce précodeur permet de pondérer les différentes voies en maximisant la capacité du canal. Ces résultats ont été obtenus sans ajout de CCE (fig. 8) :

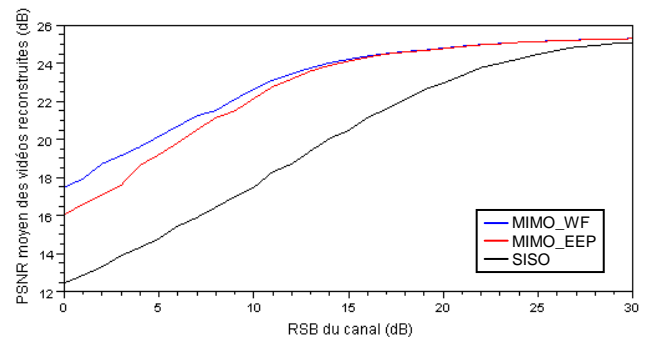


Figure 8 - Evolution du PSNR moyen en fonction du SNR pour la vidéo Foreman

Dans le cadre de la stratégie MIMO, nous utilisons quatre voies de données, le débit est donc quatre fois supérieur à la transmission SISO. L'information allouée aux différentes voies permet d'exploiter la hiérarchisation des sous-flux de données V-WTSOM en fonction de la qualité du canal de transmission. Ceci explique pourquoi les courbes MIMO donnent de meilleurs résultats que la courbe SISO. La stratégie permettant d'allouer la puissance sur les différentes voies permet d'augmenter la qualité visuelle par rapport à une transmission MIMO utilisant une pondération équivalente sur chaque voie (MIMO\_EEP).

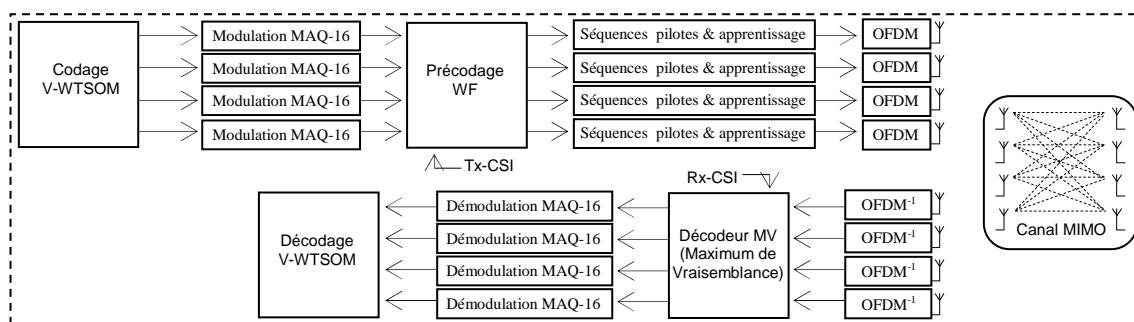


Figure 7 - La chaîne de simulation

Cette adéquation entre le contenu à transmettre et la technique de précodage permet d'assurer une qualité de service minimale dans les conditions les plus sévères du canal. Lorsque le canal est bon, les stratégies MIMO et SISO tendent vers une qualité visuelle équivalente.

L'accessibilité à une couche de donnée est définie par rapport au TEB. V-WTSOM faisant preuve d'une robustesse importante, le seuil d'accessibilité d'une couche a été fixé à  $2.10^{-1}$  de TEB (pour cet exemple). La figure 9 permet de constater l'apport en qualité de ces différentes couches de données dans le cadre de la transmission MIMO\_WF (fig. 9) :

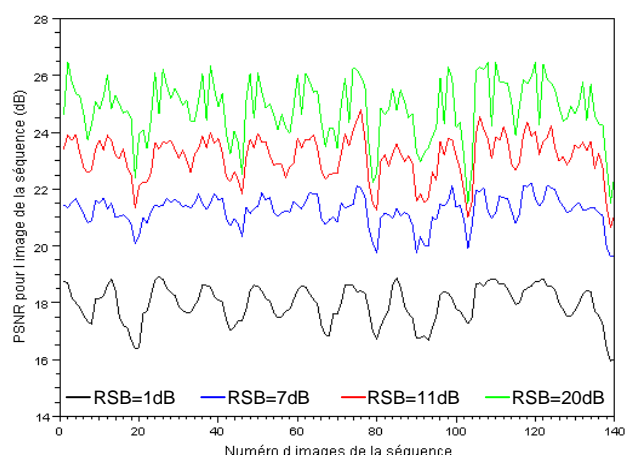


Figure 9 - PSNR des images de la vidéo Foreman suivant les couches de qualité décodées

On valide bien que la qualité de la vidéo augmente graduellement suivant la disponibilité des sous-flux à la réception. Une qualité de base est assurée même à bas débit. Puis, suivant la QoS désirée ou bien les ressources disponibles, nous pouvons envisager une qualité graduée des vidéos reconstruites. Ces résultats confirment également la robustesse dont fait preuve le codeur V-WTSOM. Malgré l'absence de CCE, nous constatons une augmentation de la qualité malgré un seuil d'acceptation des flux à  $2.10^{-1}$  de TEB.

## 5 Conclusion

Le codeur V-WTSOM a pour but de transmettre des vidéos de type vidéoconférence de manière robuste sans

ajout de CCE. Ce codeur s'appuie sur la compensation de mouvement par champs de vecteurs et sur des transformées en ondelette. Nous avons proposé une méthode robuste pour le codage des VM puis nous avons exploité le flux fortement hiérarchisé de V-WTSOM sur un canal MIMO en utilisant un précodage autorisant une stratégie d'allocation par couches de qualité. Les résultats issus des simulations ont ainsi pu mettre en avant l'intérêt d'utiliser le MIMO en tenant compte du contenu de l'information à transmettre. L'utilisation modérée de CCE ne pourra qu'améliorer les performances de V-WTSOM.

## Remerciement

Ce travail a bénéficié d'une aide de l'Agence Nationale de la Recherche portant la référence ANR-08-VERS-002.

## Références

- [1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264-AVC Video Coding Standard", in IEEE Trans. Circuits Syst. Video Technol., vol. 13, n° 7, pp. 560-576, July 2003
- [2] N. Adami, "State-of-the-art and trends in scalable video compression with wavelet-based approaches", in IEEE Trans. Circuits Syst. Video Technol., vol. 17, n°9, pp. 1238-1255, September 2007
- [3] M. Cagnazzo, T. André, M. Antonini, M. Barlaud, "A model-based motion compensated video coder with JPEG2000 compatibility", in Proc. IEEE Int. Conf. Image Process., pp. 2255-2258, Singapore, October 2004
- [4] S. Benayoune, N. Achir, K. Bousseta, K. Chen, "A study of H.264/AVC Robustness Over a Wireless Link", Signal Processing and Its Applications, pp. 1-4, Feb. 2007
- [5] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, "Error resiliency schemes in H.264/AVC standard", Journal of Visual Communication and Image Representation, vol. 17, n° 2, pp. 425-450, Apr 2006
- [6] C. Chatellier, H. Boeglen, C. Perrine, C. Olivier, O. Haeberle, "A robust joint source channel coding scheme for image transmission over the ionospheric channel", Signal Proc. Image Communication, vol. 22, n° 6, pp. 543-556, July 2007
- [7] A. Barjatya, "Block Matching Algorithm for motion estimation", in Technical Report, Utah State, 2004
- [8] W. Hamidouche, C. Perrine, Y. Pousset, C. Olivier, "Optimal solution for SVC-based video transmission over a realistic MIMO channel using precoder design", IEEE ICASSP, March 2010
- [9] IEEE Standard for Information Technology-Part 11 : Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment : Enhancements for Higher Throughput (802.11n), 2009